

A Java desktop tool for mining Moodle data

R. PEDRAZA-PEREZ, C. ROMERO AND S. VENTURA

University of Cordoba, Spain

This paper shows a Java desktop Moodle mining tool oriented not only for experts in data mining but also for newcomers like instructors and courseware authors. This tool has a wizard with a simple interface based on big friendly buttons and default options to facilitate the execution of all the steps in the data mining process.

Key Words and Phrases: Educational Data Mining, Moodle, Data Mining Tool, Wizard.

1. INTRODUCTION AND DESCRIPTION OF THE MINING TOOL

General data mining tools today range from the commercial, such as DBMiner, SPSS Clementine, SAS Enterprise Miner, IBM Intelligent Miner, to open sources, like Weka and RapidMiner. However, all these tools are not specifically designed for pedagogical/educational purposes so they are cumbersome for an educator to use because they are normally designed more for power and flexibility than for simplicity (Romero and Ventura, 2010). In order to resolve this problem, we have developed a Java desktop data-mining tool specifically for mining Moodle data. We have chosen Moodle because it is one of the most frequently used Learning Management Systems (LMS). This tool has been implemented in Java using the KEEL framework (Alcala et al., 2009) which is an open source framework for building data mining models. Our Moodle mining tool provides a wizard that has a simple interface based on big friendly buttons (see Figure 1) to facilitate its execution.



Fig. 1. Main window of the Moodle mining tool interface.

As Figure 1 shows, our wizard guides users through the most common steps involved in any knowledge discovery process, which are described below.

- *To create data files.* First of all, users must create data files starting from a Moodle database or log files provided by Moodle. Then, they can choose from among different options: creating a data file from one particular Moodle table, creating a summary data file from different Moodle tables, or creating a data file

from one Moodle log file. The summarization option allows for the integration of all the most important student information during a full course, since student and interaction data are spread over several Moodle tables. At the end of this process, a text data file is saved/created using the KEEL data format (Alcala et al., 2009).

- *To pre-process data files.* This module allows users to modify KEEL data files. The application provides several pre-processing tasks such as: discretization (to divide numerical values into categorical values), edition (to add or delete columns or rows, for example to add students' final marks) and division (to split data file into training and test data files).
- *To visualize data files.* This module will show (it is now under construction) different types of graphical representations of the aforementioned data.
- *To mine data files.* Finally, users can apply data mining techniques. Our mining tool provides a large number of algorithms grouped into three methods/tasks: classification, regression, and association. Firstly, users have to select the desired data mining method and then one of the available algorithms. Our wizard always recommends one algorithm and its parameters by default for each method in order to facilitate its usage/execution for beginners. Users can change the current selection to any other algorithm and change its parameters. Figure 2 shows the execution of the C4.5 algorithm over a summary data file and 10-fold cross-validation. Any resulting files (.tra and .test files that contain partial classification results, and the .txt file containing the model obtained) can be selected and visualized easily (see Figure 2 on the right hand side). In this example, the model obtained is a decision tree that uses IF-THEN rules and also shows a summary with the number of nodes and leaves on the tree, the number and percentage of correctly and incorrectly classified instances, etc.



Fig. 2. Some windows in the data mining wizard.

ACKNOWLEDGMENTS

This research is subsidized by projects of the Regional Government of Andalusia and the Ministry of Science and Technology, P08-TIC-3720 and TIN2008-06681-C06-03 respectively, and FEDER funds.

REFERENCES

- ALCALÁ, J., SÁNCHEZ, L., GARCÍA, S., DEL JESUS, M. ET. AL. 2009. KEEL a software tool to assess evolutionary algorithms to data mining problems. *Soft Computing*, 13(3), 307-318.
- ROMERO C, VENTURA, S. Educational Data Mining: A Review of the State-of-the-Art. 2010. *IEEE Transactions on Systems, Man and Cybernetics, part C: Applications and Reviews*, 40(6), 601-618.